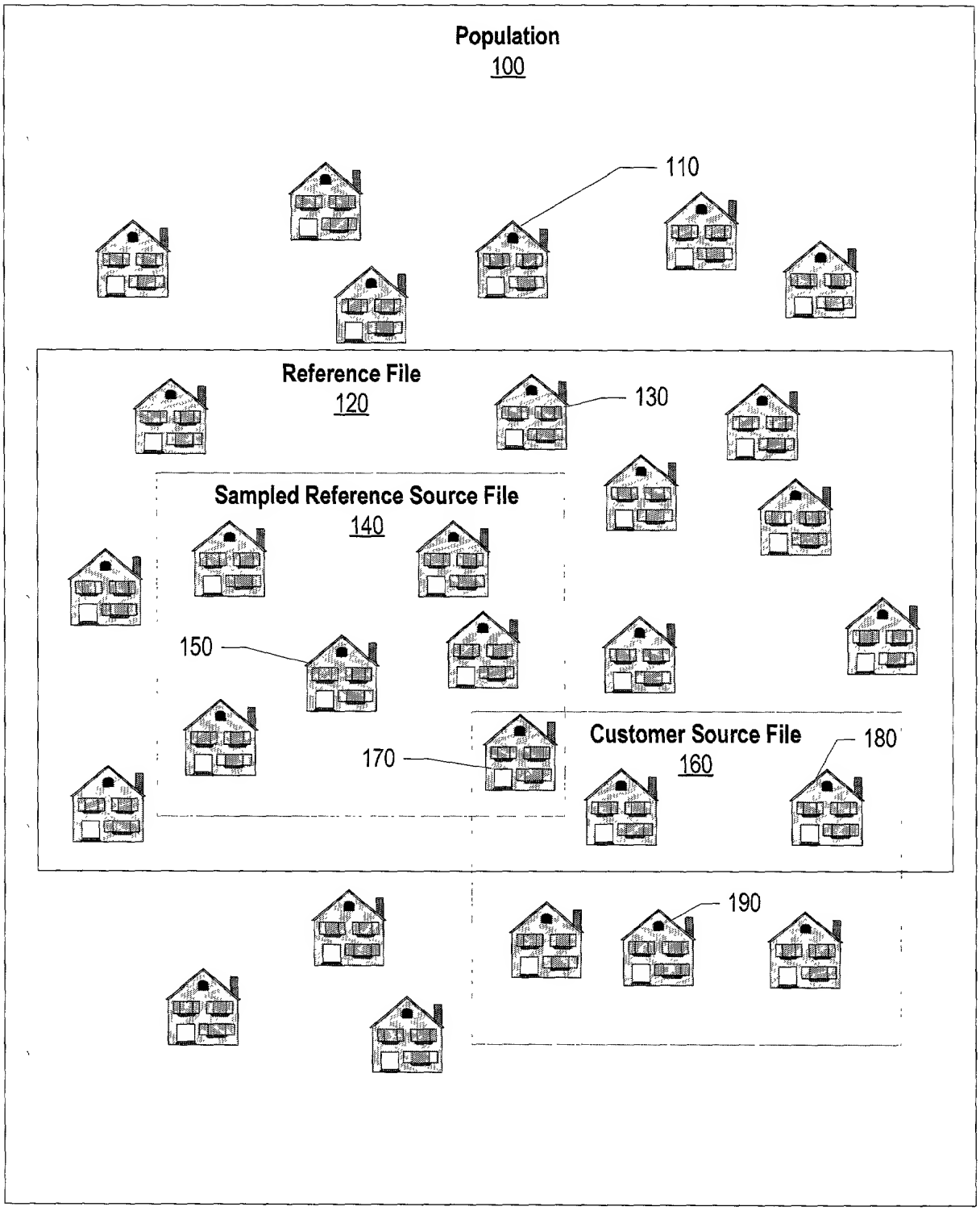
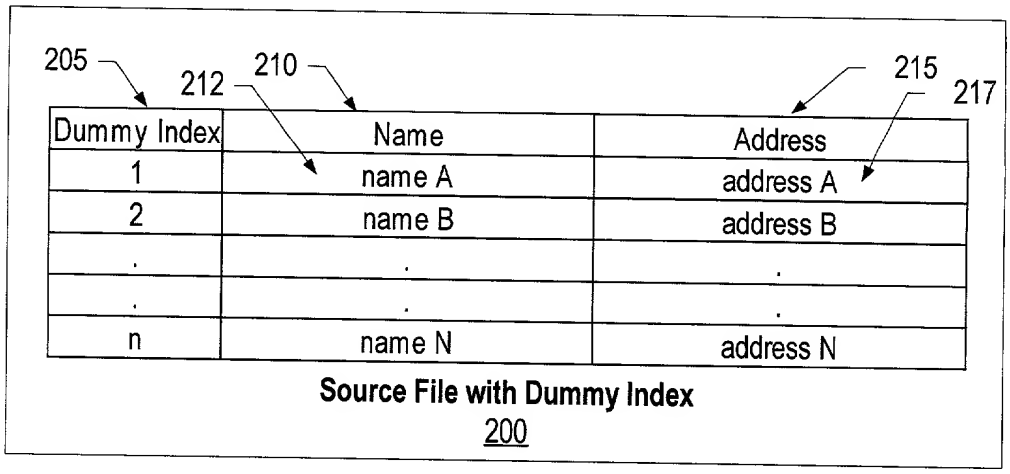
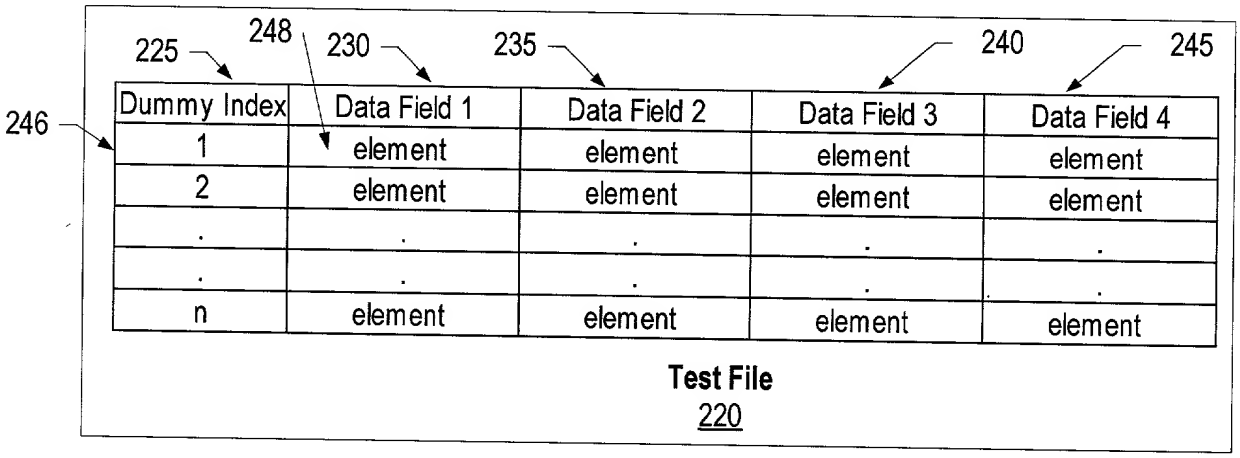

Population
100
110
Reference File
120
130
Sampled Reference Source File
140
150
Customer Source File
160
170
180
190

Figure 1


205
212
210
215
217
Dummy Index
Name
Address
1
name A
address A
2
name B
address B
.
.
.
.
.
.
n
name N
address N
Source File with Dummy Index
200


225
248
230
235
240
245
246
Dummy Index
Data Field 1
Data Field 2
Data Field 3
Data Field 4
1
element
element
element
element
2
element
element
element
element
.
.
.
.
.
.
.
.
.
.
n
element
element
element
element
Test File
220

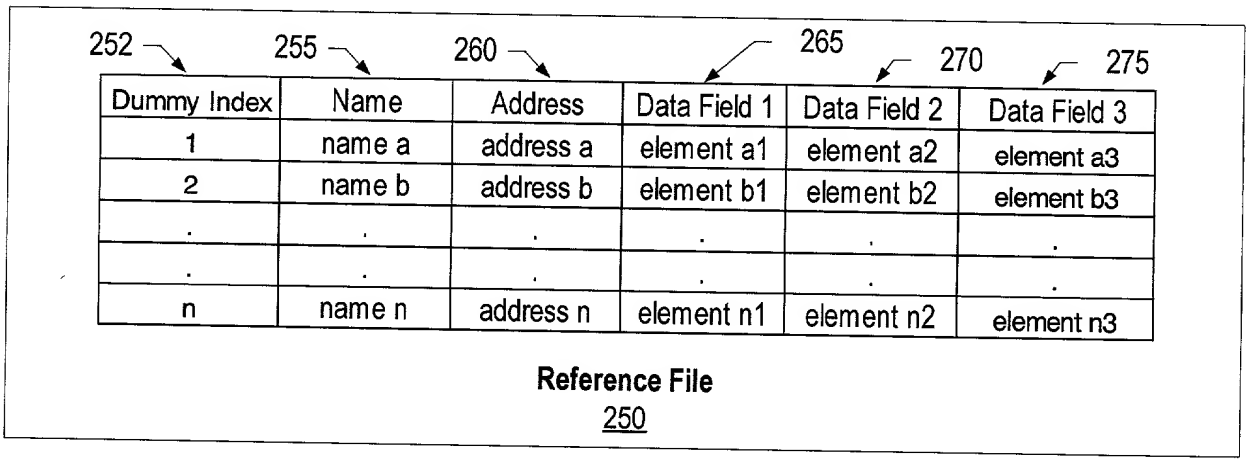

252
255
260
265
270
275
Dummy Index
Name
Address
Data Field 1
Data Field 2
Data Field 3
1
name a
address a
element a1
element a2
element a3
2
name b
address b
element b1
element b2
element b3
.
.
.
.
.
.
.
.
.
.
n
name n
address n
element n1
element n2
element n3
Reference File
250

Figure 2

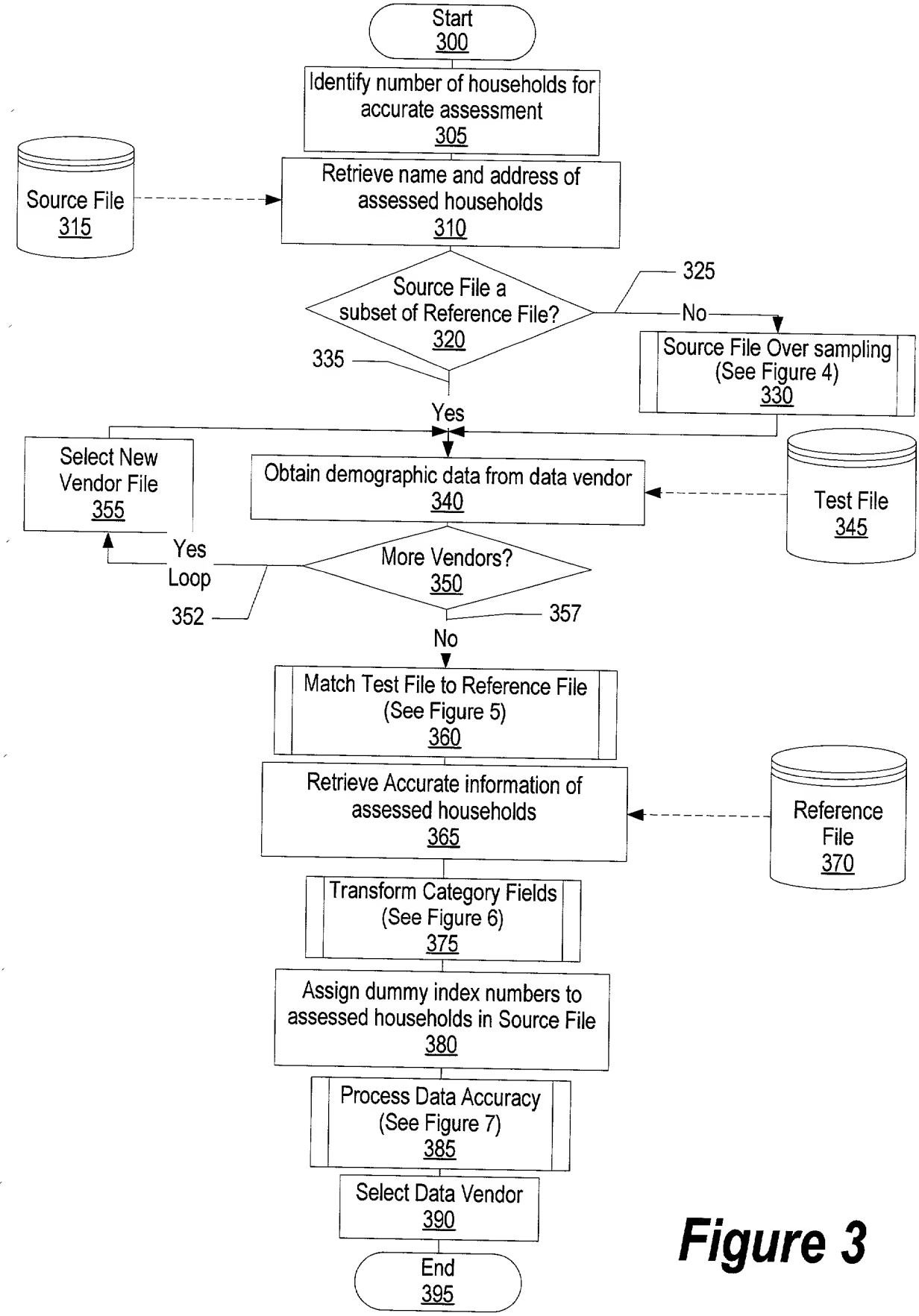

Start
300
Identify number of households for accurate assessment
305
Source File
315
Retrieve name and address of assessed households
310
Source File a subset of Reference File?
320
325
No
Source File Over sampling (See Figure 4)
330
335
Yes
Select New Vendor File
355
Obtain demographic data from data vendor
340
Test File
345
Yes
Loop
352
More Vendors?
350
357
No
Match Test File to Reference File (See Figure 5)
360
Retrieve Accurate information of assessed households
365
Reference File
370
Transform Category Fields (See Figure 6)
375
Assign dummy index numbers to assessed households in Source File
380
Process Data Accuracy (See Figure 7)
385
Select Data Vendor
390
End
395

Figure 3

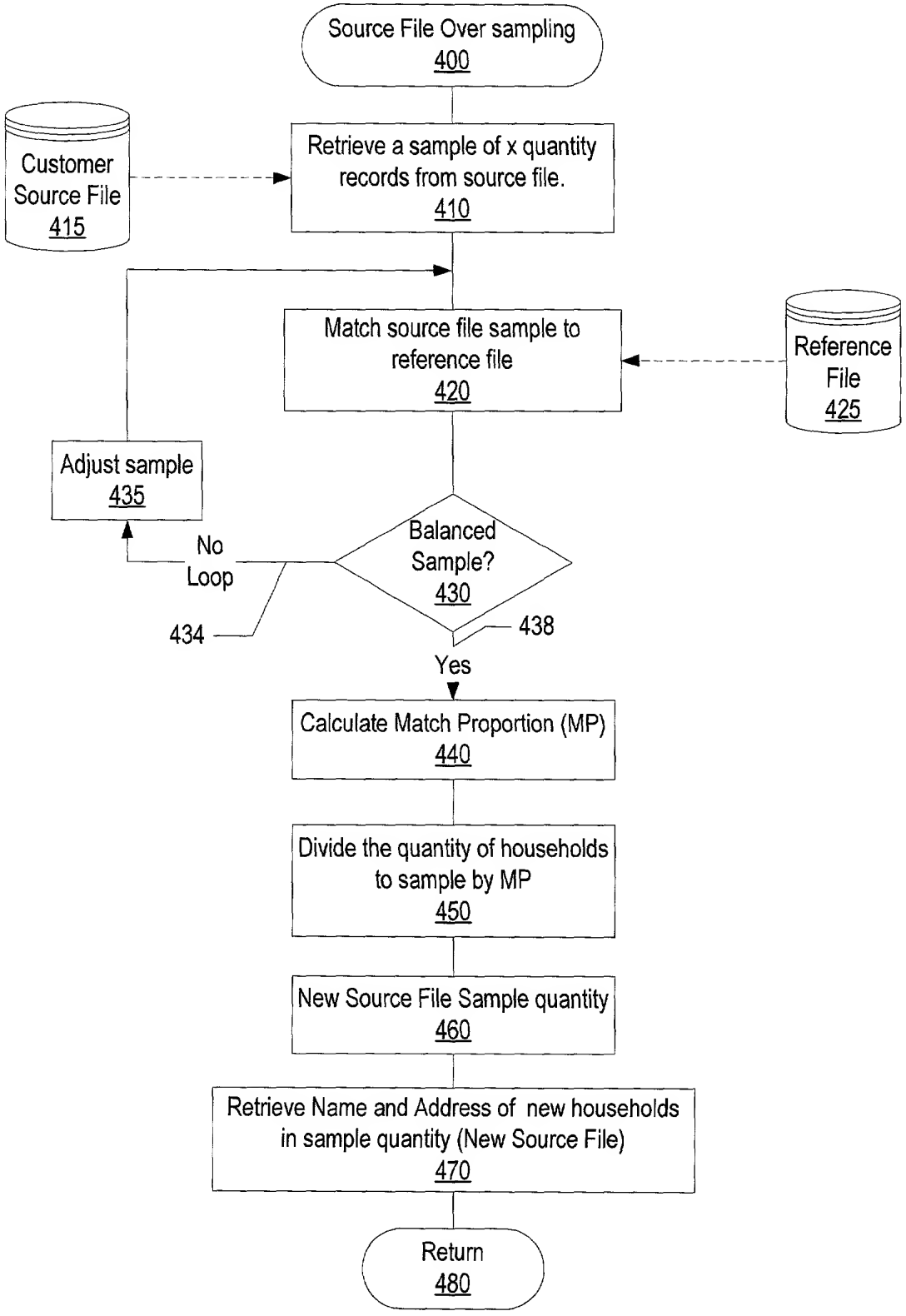

Source File Over sampling
400
Customer Source File
415
Retrieve a sample of x quantity records from source file.
410
Match source file sample to reference file
420
Reference File
425
Adjust sample
435
No Loop
434
Balanced Sample?
430
438
Yes
Calculate Match Proportion (MP)
440
Divide the quantity of households to sample by MP
450
New Source File Sample quantity
460
Retrieve Name and Address of new households in sample quantity (New Source File)
470
Return
480

Figure 4

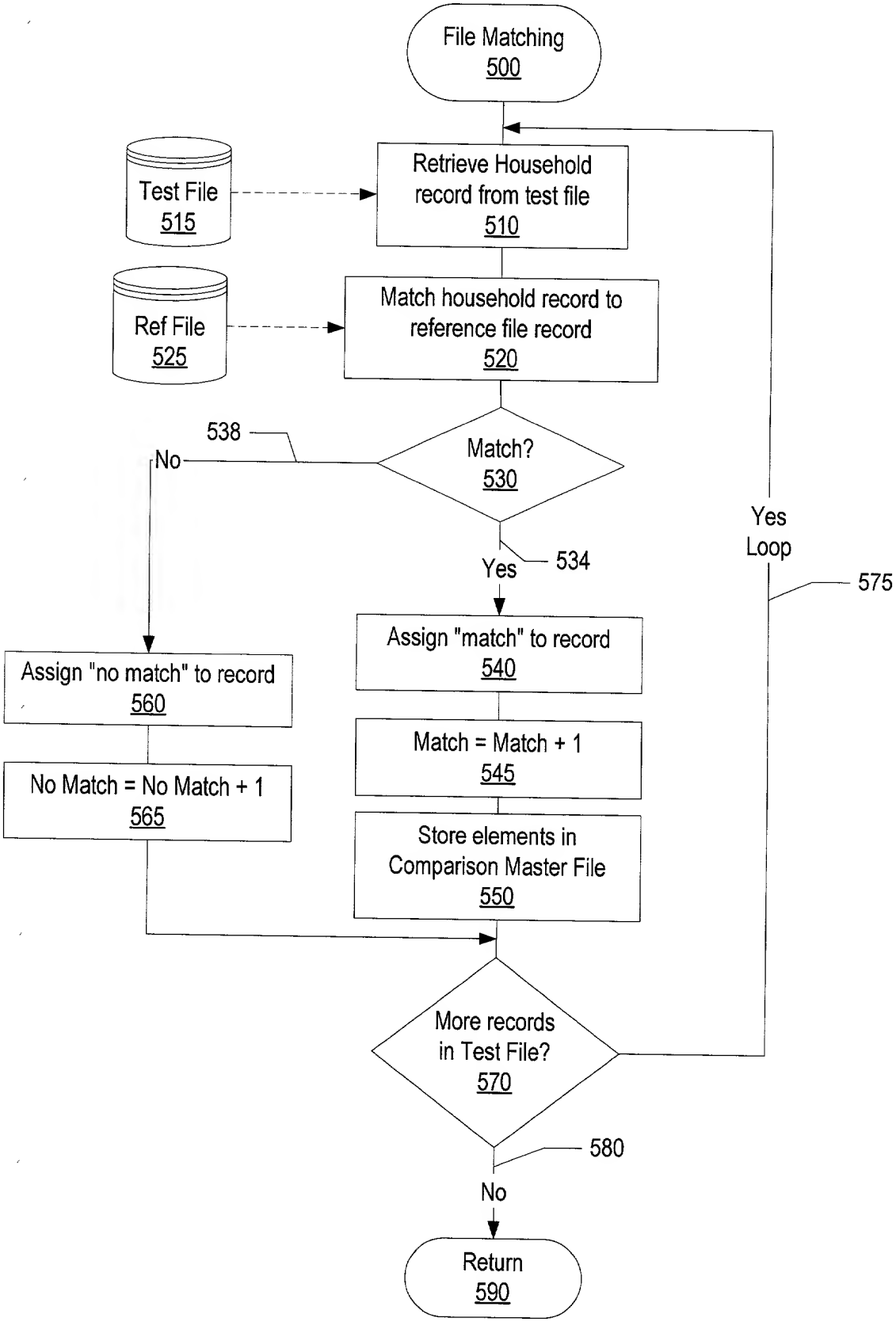

File Matching
500
Test File
515
Retrieve Household record from test file
510
Ref File
525
Match household record to reference file record
520
538
No
Match?
530
Yes
534
Yes Loop
575
Assign "match" to record
540
Match = Match + 1
545
Store elements in Comparison Master File
550
Assign "no match" to record
560
No Match = No Match + 1
565
More records in Test File?
570
580
No
Return
590

Figure 5

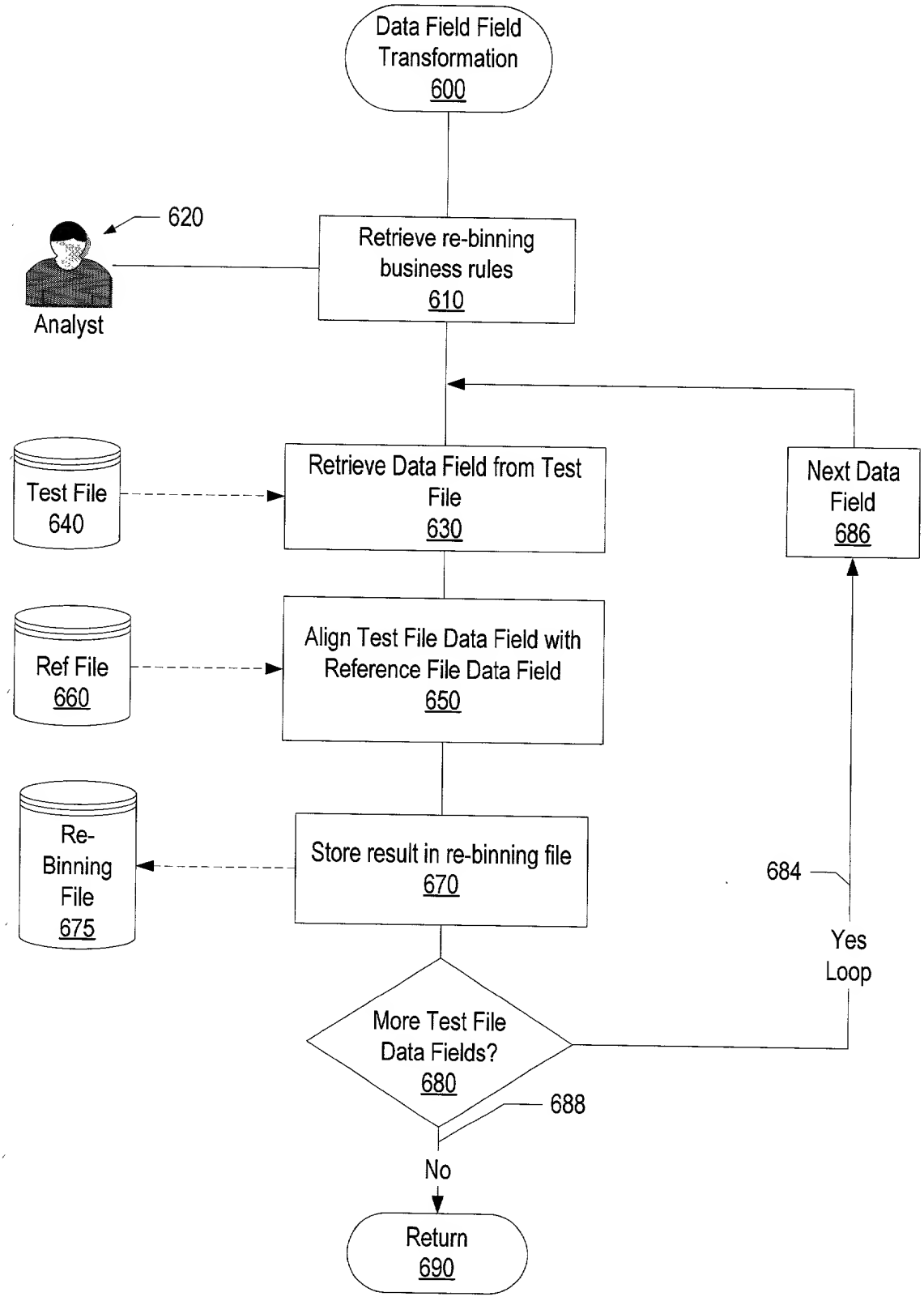

Data Field Field Transformation 600
620
Analyst
Retrieve re-binning business rules 610
Test File 640
Retrieve Data Field from Test File 630
Next Data Field 686
Ref File 660
Align Test File Data Field with Reference File Data Field 650
Re-Binning File 675
Store result in re-binning file 670
684
Yes Loop
More Test File Data Fields? 680
688
No
Return 690

Figure 6

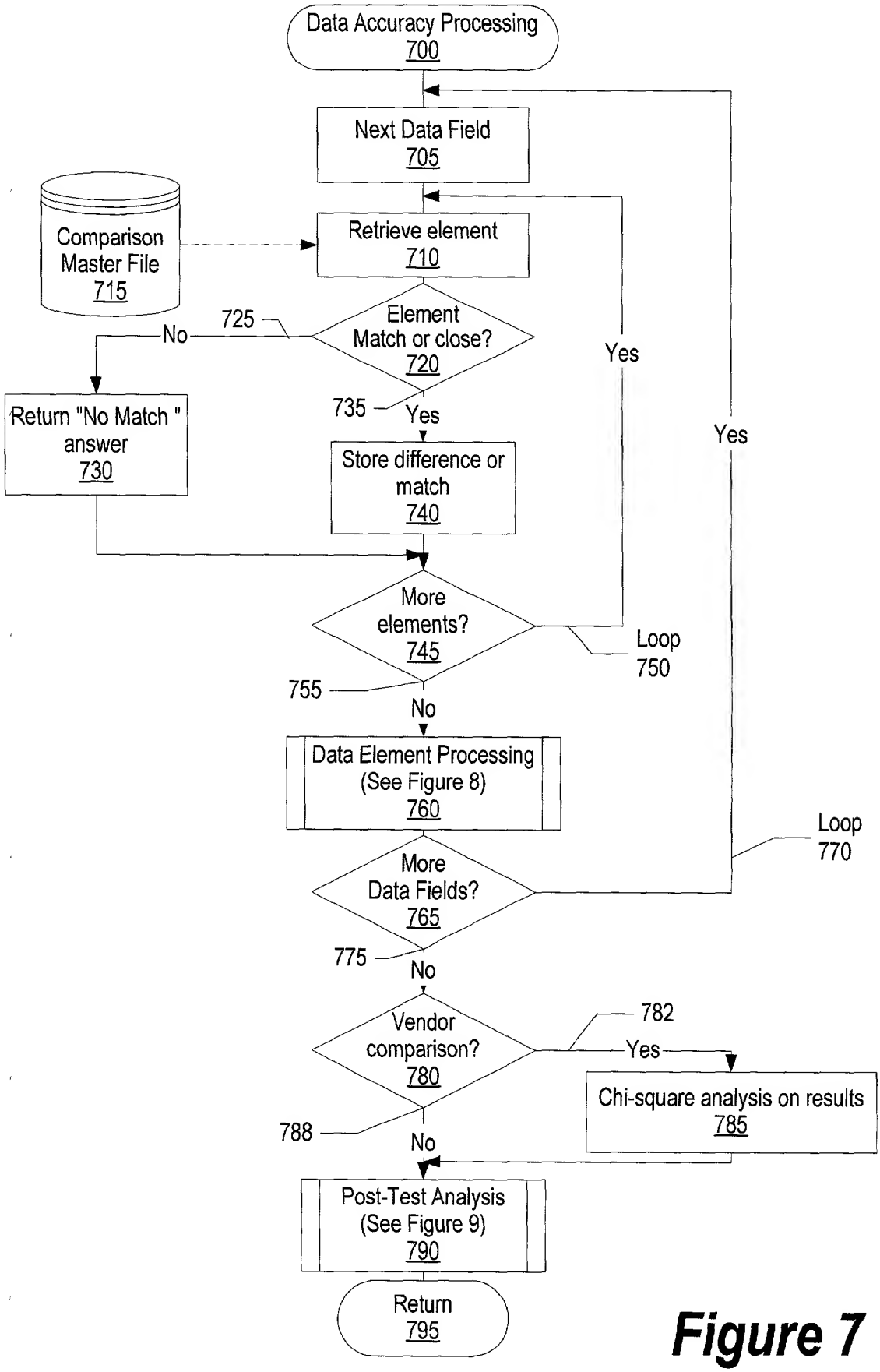

Data Accuracy Processing
700
Next Data Field
705
Comparison Master File
715
Retrieve element
710
Element Match or close?
720
725
No
Yes
Return "No Match " answer
730
735
Yes
Store difference or match
740
More elements?
745
Loop
750
Yes
755
No
Data Element Processing
(See Figure 8)
760
Loop
770
More Data Fields?
765
775
No
Vendor comparison?
780
782
Yes
Chi-square analysis on results
785
788
No
Post-Test Analysis
(See Figure 9)
790
Return
795

Figure 7

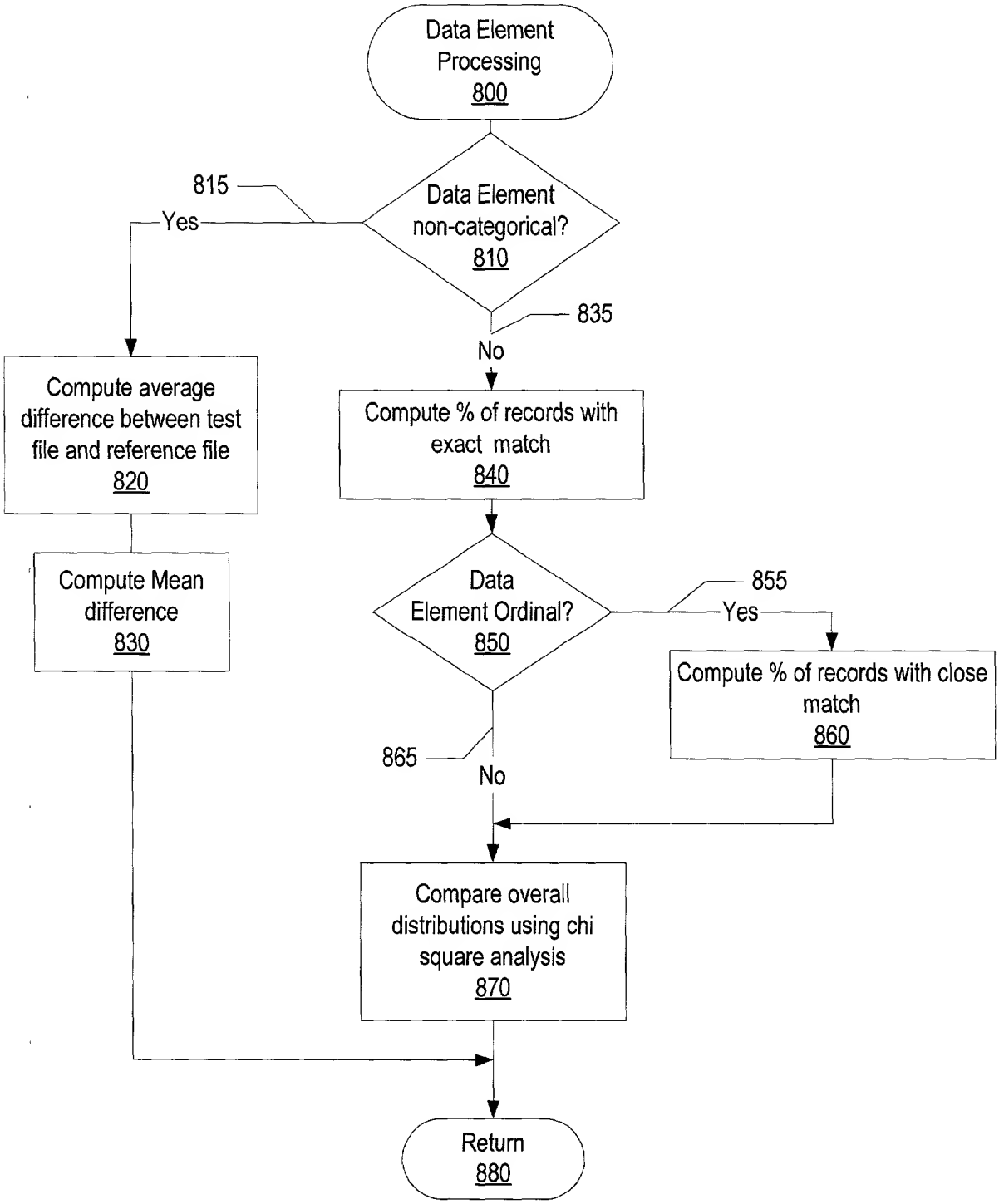

Data Element Processing
800
Data Element non-categorical?
810
815
Yes
835
No
Compute average difference between test file and reference file
820
Compute % of records with exact match
840
Compute Mean difference
830
Data Element Ordinal?
850
855
Yes
Compute % of records with close match
860
865
No
Compare overall distributions using chi square analysis
870
Return
880

Figure 8

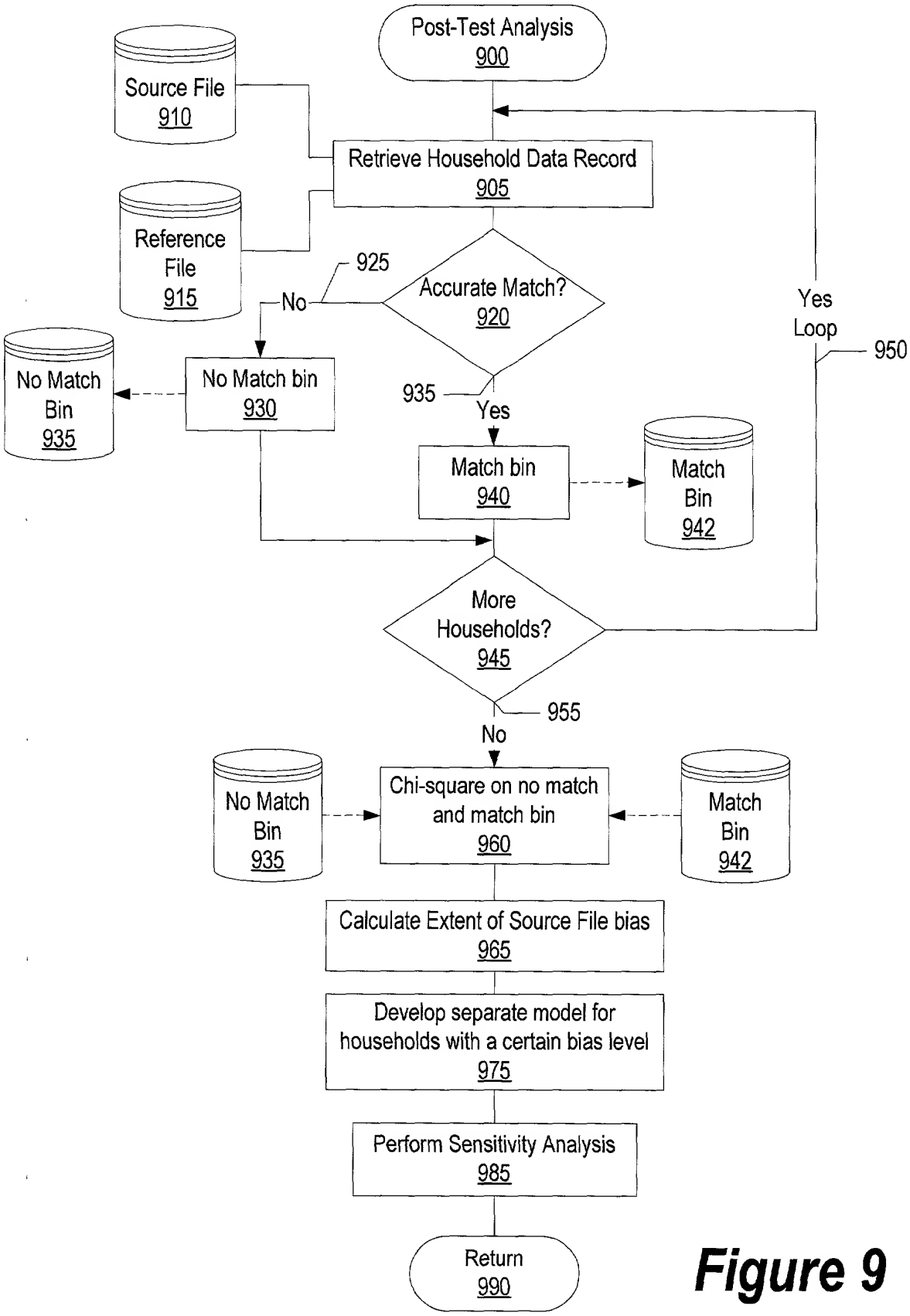

Post-Test Analysis
900
Source File
910
Retrieve Household Data Record
905
Reference File
915
925
No
Accurate Match?
920
Yes
Loop
950
No Match Bin
935
No Match bin
930
935
Yes
Match bin
940
Match Bin
942
More Households?
945
955
No
No Match Bin
935
Chi-square on no match and match bin
960
Match Bin
942
Calculate Extent of Source File bias
965
Develop separate model for households with a certain bias level
975
Perform Sensitivity Analysis
985
Return
990

Figure 9

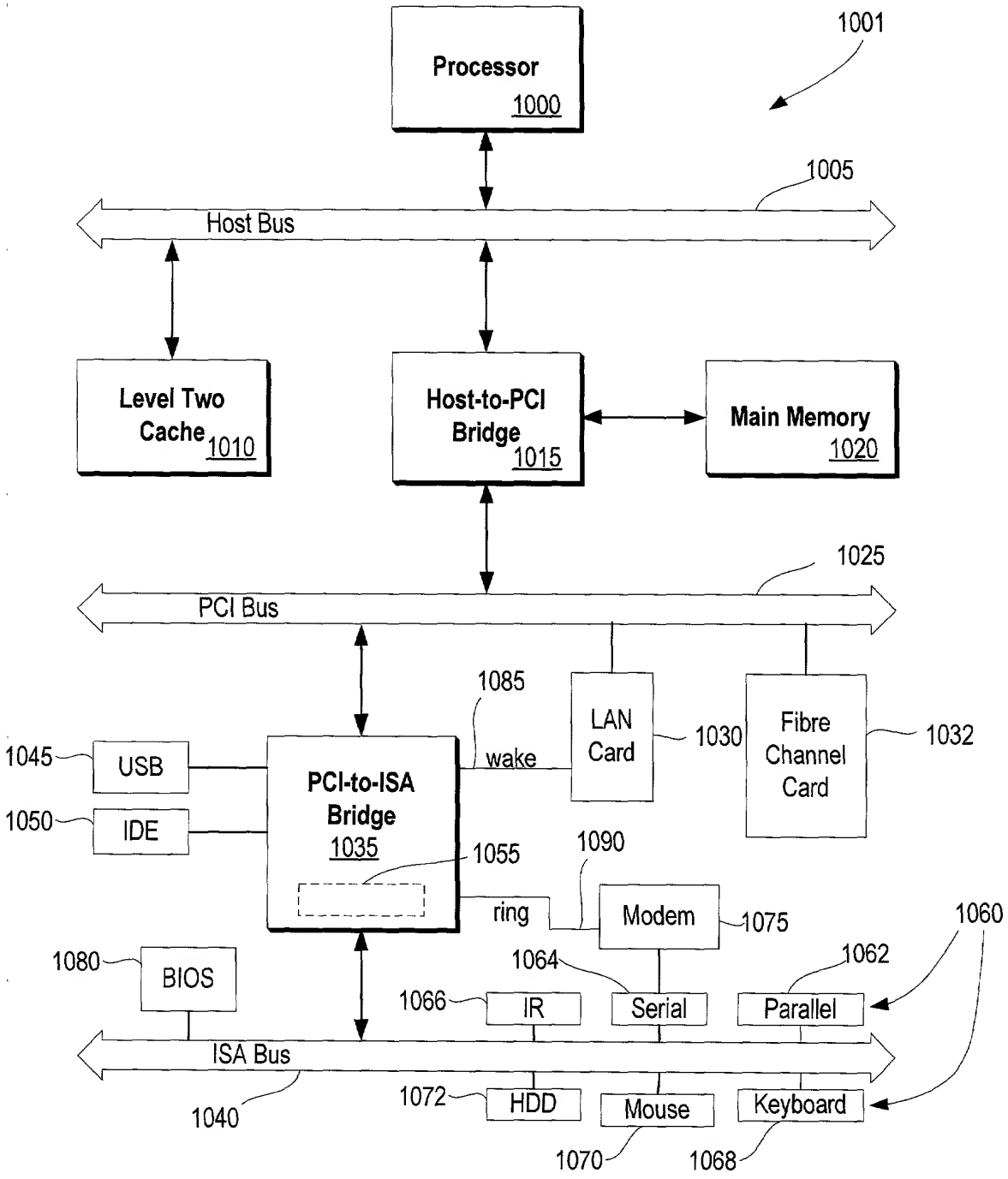

1001
Processor
1000
1005
Host Bus
Level Two
Cache
1010
Host-to-PCI
Bridge
1015
Main Memory
1020
1025
PCI Bus
1085
wake
LAN
Card
1030
Fibre
Channel
Card
1032
1045
USB
1050
IDE
PCI-to-ISA
Bridge
1035
1055
1090
ring
Modem
1075
1060
1080
BIOS
1064
1062
1066
IR
Serial
Parallel
ISA Bus
1040
1072
HDD
Mouse
Keyboard
1070
1068

Figure 10